CLAIMS

What is claimed is:

1. A method for identifying a collection of polymorphisms from nucleic acid molecules in a sample by analyzing a subset of the molecules, comprising the steps of:

   a. obtaining a nucleic acid-containing sample;

   b. treating the nucleic acid molecules in said sample to produce a reduced representation of nucleic acid fragments selected in a sequence-dependent manner by a method comprising:

      i. fractionating said nucleic acid molecules to produce nucleic acid fragments; and

      ii. selecting a subset of said nucleic acid fragments,

      wherein either (i) or (ii) or both (i) and (ii) are performed in a sequence-dependent manner;

   c. analyzing the reduced representation to identify pairs of fragments corresponding to the same chromosomal location, wherein fragments corresponding to the same chromosomal location are orthologous sequences; and

   d. comparing pairs of orthologous sequences to identify polymorphisms between said sequences.

2. The method of Claim 1, wherein the polymorphisms are single nucleotide polymorphisms.

3. The method of Claim 1, wherein the nucleic acid-containing sample is pooled from more than one individual.

4. The method of Claim 1, wherein the nucleic acid molecules are DNA.

5. The method of Claim 1, wherein the nucleic acid molecules are RNA.

sub a3 > 6. The method of Claim 3, wherein the individuals share a particular trait.

7. The method of Claim 6, where the trait is a disorder.

sub a4 > 8. The method of Claim 1, wherein step (b)(i) is performed by one or more restriction endonucleases.

9. The method of Claim 8, wherein the one or more restriction endonucleases are selected from the group consisting of *Bgl*II, *Xho*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Pst*I, and *Hae*III.

10. The method of Claim 1, wherein step (b)(ii) is performed using an agarose gel.

11. The method of Claim 1, wherein step (b)(ii) is performed using high pressure liquid chromatography (HPLC).

Sub C2 > 12. The method of Claim 1, wherein step (b)(ii) is performed by selecting nucleic acid fragments which hybridize to selected additional nucleic acid sequences.

13. The method of Claim 1, wherein step (c) and/or step (d) are performed by determining at least a portion of the nucleic acid sequence of the orthologous sequences.

14. A method for identifying a collection of polymorphisms from nucleic acid molecules in a sample by analyzing a subset of the molecules, comprising the steps of:

a. obtaining a nucleic acid-containing sample to be assessed;

b. treating nucleic acid molecules in said sample to produce a reduced representation of nucleic acid fragments selected in a sequence-dependent manner by a method comprising:

i. fractionating said nucleic acid molecules with one or more restriction endonucleases to produce nucleic acid fragments; and

ii. selecting a subset of said nucleic acid fragments using size fractionation;

wherein either (i) or (ii) or both (i) and (ii) are performed in a sequence-dependent manner;

c. analyzing the reduced representation to identify pairs of fragments corresponding to the same chromosomal location, wherein fragments corresponding to the same chromosomal location are orthologous sequences; and

d. comparing pairs of orthologous sequences to identify polymorphisms between said orthologous sequences,

thereby identifying a collection of polymorphisms from said nucleic acid molecules.

15. The method of Claim 14, wherein the polymorphisms are single nucleotide polymorphisms.

16. The method of Claim 14, wherein the nucleic acid-containing sample is pooled from more than one individual.

17. The method of Claim 14, wherein the nucleic acid molecules are DNA.

18. The method of Claim 14, wherein the nucleic acid molecules are RNA.

19. The method of Claim 16, wherein the individuals share a particular trait.

20. The method of Claim 19, wherein the trait is a disorder.

21. The method of Claim 14, wherein the one or more restriction endonucleases are selected from the group consisting of *Bgl*II, *Xho*I, *Eco*RI, *Eco*RV, *Hin*dIII, *Pst*I, and *Hae*III.

22. The method of Claim 14, wherein step (b)(ii) is performed using an agarose gel.

23. The method of Claim 14, wherein step (b)(ii) is performed using high pressure liquid chromatography (HPLC).

24. The method of Claim 14, wherein step (b)(ii) is performed by selecting nucleic acid fragments which hybridize to selected additional nucleic acid sequences.

25. The method of Claim 14, wherein step (c) and/or step (d) are performed by determining at least a portion of the nucleic acid sequence of the orthologous sequences.

26. The method of Claim 14, wherein the one or more restriction endonucleases cleave DNA on average about once every 2000 base pairs.

27. The method of Claim 14, wherein the subset of (b)(ii) is in a size range selected from the group consisting of: from about 380 base pairs to about 480 base pairs, from about 400 base pairs to about 500 base pairs, from about 480 base pairs to

about 580 base pairs, from about 500 base pairs to about 600 base pairs, and from about 540 base pairs to about 640 base pairs.

28. A method for genotyping a nucleic acid sample for polymorphisms in nucleic acid fragments contained in a reduced representation, comprising the steps of:

a. obtaining a nucleic acid-containing sample;

b. treating the nucleic acid molecules in said sample to produce a reduced representation of nucleic acid fragments selected in a sequence-dependent manner by a method comprising:

i. fractionating said nucleic acid molecules to produce nucleic acid fragments; and

ii. selecting a subset of said nucleic acid fragments,

wherein either (i) or (ii) or both (i) and (ii) are performed in a sequence-dependent manner; and

c. analyzing the nucleic acid fragments contained in the reduced representation to assess the genotype at one or more polymorphic sites.

29. The method of Claim 28, wherein step (b)(ii) is performed using an agarose gel.

30. The method of Claim 28, wherein step (b)(ii) is performed using high pressure liquid chromatography (HPLC).

31. The method of Claim 28, wherein step (b)(ii) is performed by selecting nucleic acid fragments which hybridize to selected additional nucleic acid sequences.

32. The method of Claim 28, wherein step (c) is performed by determining at least a portion of the nucleic acid sequence of the nucleic acid fragments.

Sub a8

33. The method of Claim 28, wherein step (c) is performed by attaching specific oligonucleotide linker sequences to the fragments in the reduced representation and then amplifying said fragments.

34. The method of Claim 33, wherein the amplification is performed by polymerase chain reaction using primers complementary to the linker sequences.

35. The method of Claim 33, wherein the amplification is performed by cloning the fragments in an organism.

36. The method of Claim 28, wherein step (c) is performed by performing single-base extension reactions on the reduced representation.

37. The method of Claim 33, wherein step (c) is performed by performing single-base extension reactions on the reduced representation.

38. The method of Claim 28, wherein step (c) is performed by hybridization to an oligonucleotide array.

39. The method of Claim 33, wherein step (c) is performed by hybridization to an oligonucleotide array.

40. The method of Claim 28, wherein step (c) is performed by an oligo ligation assay.

41. The method of Claim 33, wherein step (c) is performed by an oligo ligation assay.

42. The method of Claim 1, wherein step (c) is performed by the following steps:

a. comparing the sequences of the two members of a proposed pair, wherein the two sequences are further analyzed if the two sequences are at least 80% identical over at least 80% of the length of the shorter of the two sequences;

b. aligning the two sequences identified from (a), wherein the two sequences are further analyzed if the two sequences are identical over 10 or more bases within the first 50 bases and the last 50 bases of the sequences;

c. identifying candidate single nucleotide polymorphisms in the sequences of (b), wherein the two sequences are further analyzed if the number of candidate single nucleotide polymorphisms does not exceed 1% of the total number of bases in the shorter of the two sequences, wherein two sequences which meet the criteria of (a) - (c) qualify as a candidate match;

d. repeating (a) - (c) for all proposed pairs; and

e. determining the number of candidate matches for the same chromosomal location, wherein said candidate matches are accepted if said number of matches does not exceed expectations,

wherein accepted candidate matches are considered a pair.

43. The method of Claim 42, wherein said expectations are determined according to binomial or Poisson distributions.

44. The method of Claim 14, wherein step (c) is performed by the following steps:

a. comparing the sequences of the two members of a proposed pair, wherein the two sequences are further analyzed if the two sequences are at least 80% identical over at least 80% of the length of the shorter of the two sequences;

b. aligning the two sequences identified from (a), wherein the two sequences are further analyzed if the two sequences are identical over 10 or more bases within the first 50 bases or the last 50 bases of the sequences;

c. identifying candidate single nucleotide polymorphisms in the sequences of (b), wherein the two sequences are further analyzed if the number of candidate single nucleotide polymorphisms does not exceed 1% of the total number of bases in the shorter of the two sequences, wherein two sequences which meet the criteria of (a) - (c) qualify as a candidate match;

d. repeating (a) - (c) for all proposed pairs; and

e. determining the number of candidate matches for the same chromosomal location, wherein said candidate matches are accepted if said number of matches does not exceed expectations,

wherein accepted candidate matches are considered a pair.

45. The method of Claim 44, wherein said expectations are determined according to binomial or Poisson distributions.

46. A method for determining a limited population of polymorphisms from nucleic acid molecules in a sample, comprising the steps of:

a. obtaining a nucleic acid-containing sample to be assessed;

b. treating nucleic acid molecules in said sample to produce nucleic acid fragments selected in a sequence-dependent manner by a method comprising:

i. fractionating said nucleic acid molecules to produce nucleic acid fragments; and

ii. selecting a subset of said nucleic acid fragments;

wherein either (i) or (ii) or both (i) and (ii) are done in a sequence-dependent manner;

c. selecting from said subset nucleic acid fragments which occur at a corresponding chromosomal locus, thereby producing a pair, and

d. identifying polymorphisms between fragments of a pair;

thereby determining a limited population of polymorphisms from said nucleic acid-containing sample.

47. A method for determining a limited population of polymorphisms from nucleic acid molecules in a sample, comprising the steps of:

a. obtaining a nucleic acid-containing sample to be assessed;

b. treating nucleic acid molecules in said sample to produce nucleic acid fragments selected in a sequence-dependent manner by a method comprising:

i. fractionating said nucleic acid molecules with one or more restriction endonucleases to produce nucleic acid fragments; and

ii. selecting a subset of said nucleic acid fragments using size fractionation;

wherein either (i) or (ii) or both (i) and (ii) are done in a sequence-dependent manner;

c. selecting from said subset nucleic acid fragments which occur at a corresponding chromosomal locus, thereby producing a pair, and

d. identifying polymorphisms between fragments of a pair;

thereby determining a limited population of polymorphisms from said nucleic acid-containing sample.

48. A method for genotyping a nucleic acid-containing sample from an individual for polymorphisms, the method comprising:

a. obtaining a first nucleic acid-containing sample to be assessed;

b. treating nucleic acid molecules in said sample to produce a reduced representation of nucleic acid fragments selected in a sequence-dependent manner by a method comprising:

i. fractionating said nucleic acid molecules to produce nucleic acid fragments; and

ii. selecting a subset of said nucleic acid fragments;

wherein either (i) or (ii) or both (i) and (ii) are done in a sequence-dependent manner;

c. analyzing the reduced representation to identify pairs of fragments corresponding to the same chromosomal location, wherein fragments corresponding to the same chromosomal location are orthologous sequences;

d. comparing pairs of orthologous sequences to identify polymorphisms between the orthologous sequences;

e. obtaining a second nucleic acid-containing sample from an individual to be assessed; and

f. analyzing said second nucleic acid-containing sample to assess the genotype at one or more polymorphisms identified in (d).

49. A method according to Claim 48, wherein the second nucleic acid-containing sample is treated by a method identical to step (b).

Add